

Small Superposition Dimension and Active Set Construction for Multivariate Integration Under Modest Error Demand

A. D. Gilbert and G. W. Wasilkowski

March 6, 2017

Abstract: Constructing *active sets* is a key part of the Multivariate Decomposition Method. An algorithm for constructing optimal or quasi-optimal active sets is proposed in the paper. By numerical experiments, it is shown that the new method can provide sets that are significantly smaller than the sets constructed by the already existing method. The experiments also show that the *superposition dimension* could surprisingly be very small, at most 3, when the error demand is not smaller than 10^{-3} and the weights decay sufficiently fast.

1 Introduction

In this short paper, we consider approximating integrals with infinitely many variables. We focus on approximations with a modest error demand, aiming at problems in, e.g., *Mathematical Finance* and *Uncertainty Quantification*, where the underlying stochastic process is not known and hence only rough approximations are needed. In our tests we use $\varepsilon = 10^{-n}$ for $n = 1, 2, 3$ as the error demands.

The functions to be integrated belong to $\{\gamma_{\mathbf{u}}\}_{\mathbf{u} \in \mathbb{N}_+}$ -weighted tensor product Banach spaces \mathcal{F}_{γ} which allow for the decomposition

$$f(\mathbf{x}) = \sum_{\mathbf{u} \in \mathbb{N}_+, |\mathbf{u}| < \infty} f_{\mathbf{u}}(\mathbf{x}).$$

Here the summation is with respect to finite subsets \mathbf{u} of positive integers and each $f_{\mathbf{u}}$ depends only on the variables x_j with $j \in \mathbf{u}$. We also assume that the weights $\gamma_{\mathbf{u}}$ have a product form.

Integrals of such functions can be approximated by the *Multivariate Decomposition Method*, which is a refined version of the *Changing Dimension Algorithm* introduced in [4]. An essential part of those methods is the construction of an *active set* $\mathfrak{U}(\varepsilon)$ of subsets

\mathbf{u} such that the integral of $\sum_{\mathbf{u} \notin \mathfrak{U}(\varepsilon)} f_{\mathbf{u}}$ can be neglected since it is bounded from above by

$$\varepsilon \left\| \sum_{\mathbf{u} \notin \mathfrak{U}(\varepsilon)} f_{\mathbf{u}} \right\|_{\mathcal{F}_{\gamma}} \quad \text{for all } f \in \mathcal{F}_{\gamma}.$$

In other words, it is enough to approximate integrals of the partial sum

$$\sum_{\mathbf{u} \in \mathfrak{U}(\varepsilon)} f_{\mathbf{u}}.$$

We would like to construct possibly small active sets and such that the largest cardinality among its elements \mathbf{u} ,

$$d(\mathfrak{U}(\varepsilon)) := \max_{\mathbf{u} \in \mathfrak{U}(\varepsilon)} |\mathbf{u}|,$$

is also small. This is because the partial sum $\sum_{\mathbf{u} \in \mathfrak{U}(\varepsilon)} f_{\mathbf{u}}$, that can be considered instead of the infinite sum $\sum_{\mathbf{u} \in \mathbb{N}_+, |\mathbf{u}| < \infty}$, has a small number $|\mathfrak{U}(\varepsilon)|$ of functions $f_{\mathbf{u}}$, each depending on no more than $d(\mathfrak{U}(\varepsilon))$ variables.

A specific construction of such sets (denoted by $\mathfrak{U}^{\text{PW}}(\varepsilon)$) was proposed in [7] and it was shown there that the largest cardinality among all $\mathbf{u} \in \mathfrak{U}^{\text{PW}}(\varepsilon)$ grows very slowly with decreasing ε ,

$$d(\mathfrak{U}^{\text{PW}}(\varepsilon)) = O\left(\frac{\ln(1/\varepsilon)}{\ln(\ln(1/\varepsilon))}\right) \quad \text{as } \varepsilon \rightarrow 0.$$

Moreover the size $|\mathfrak{U}^{\text{PW}}(\varepsilon)|$ grows polynomially in $1/\varepsilon$. However, the asymptotic constants in the big- O notation were not investigated and, as we shall see, they could be very large.

This is why in this paper we consider constructing possibly smallest active sets denoted by $\mathfrak{U}^{\text{opt}}(\varepsilon)$. As we will show by examples, the difference between the size of $\mathfrak{U}^{\text{opt}}(\varepsilon)$ and $\mathfrak{U}^{\text{PW}}(\varepsilon)$ could be very large. We also provide a construction of *quasi-optimal* sets, denoted by $\mathfrak{U}^{\text{q-opt}}(\varepsilon)$, which sometimes are only slightly larger than the optimal $\mathfrak{U}^{\text{opt}}(\varepsilon)$; however, their construction is less expensive.

We are also interested in active sets with the smallest $d(\mathfrak{U}(\varepsilon))$. This leads to the following concept of ε -*superposition dimension* (or *superposition dimension* for short) defined by

$$d^{\text{sup}}(\varepsilon) := \min\{d(\mathfrak{U}(\varepsilon)) : \mathfrak{U}(\varepsilon) \text{ is an active set}\}.$$

Since the optimal active sets in our experiments have very small $d(\mathfrak{U}^{\text{opt}}(\varepsilon))$, this implies that the superposition dimension is also small.

Note that our concept of the superposition dimension depends on the integration problem as well as the error demand ε . Hence it is in the same spirit as the definition of *truncation dimension* introduced recently in [3]. They are different from the definitions in statistical literature, see, e.g., [1, 5, 6, 9], where superposition and truncation dimensions are defined based on ANOVA decompositions and without any relation to the integration problem or the error demand ε . Moreover, the dimensions from [1, 5, 9] depend on specific functions, whereas the dimensions in [3] and in this paper are defined in the worst case sense, i.e., are relevant to all functions from the space \mathcal{F}_{γ} .

Although the algorithms for constructing $\mathfrak{U}^{\text{q-opt}}$ and $\mathfrak{U}^{\text{opt}}$ work for rather general problems and spaces, we applied them to the integration problem and for weighted spaces of functions with mixed first order partial derivatives bounded in L_p norms for $p \in [1, \infty]$. Such spaces have often been considered (mostly for $p = 2$) when dealing with *quasi-Monte Carlo methods*.

The results depend on how fast the weights converge to zero. In the experiments, we considered

$$\gamma_{\mathbf{u}} = \prod_{j \in \mathbf{u}} j^{-a} \quad \text{for } a = 2, 3, 4.$$

For $p = 1$, the construction from [7] is optimal, and it yields the following results:

$$\begin{aligned} d^{\text{sup}}(10^{-1}) &= \begin{cases} 1 & \text{for } a = 4, \\ 2 & \text{for } a = 3, \\ 2 & \text{for } a = 2, \end{cases} \quad \text{and} \quad |\mathfrak{U}^{\text{PW}}(10^{-1})| = \begin{cases} 2 & \text{for } a = 4, \\ 4 & \text{for } a = 3, \\ 6 & \text{for } a = 2. \end{cases} \\ d^{\text{sup}}(10^{-2}) &= \begin{cases} 2 & \text{for } a = 4, \\ 2 & \text{for } a = 3, \\ 3 & \text{for } a = 2 \end{cases} \quad \text{and} \quad |\mathfrak{U}^{\text{PW}}(10^{-2})| = \begin{cases} 6 & \text{for } a = 4, \\ 8 & \text{for } a = 3, \\ 22 & \text{for } a = 2. \end{cases} \\ d^{\text{sup}}(10^{-3}) &= \begin{cases} 2 & \text{for } a = 4, \\ 3 & \text{for } a = 3, \\ 4 & \text{for } a = 2 \end{cases} \quad \text{and} \quad |\mathfrak{U}^{\text{PW}}(10^{-3})| = \begin{cases} 10 & \text{for } a = 4, \\ 22 & \text{for } a = 3, \\ 114 & \text{for } a = 2. \end{cases} \end{aligned}$$

For $p > 1$, \mathfrak{U}^{PW} are no longer optimal; however they are not much worse than optimal sets when p is relatively close to 1. Moreover, for all the tests we have performed $d(\mathfrak{U}^{\text{PW}}(\varepsilon))$ is very close to the superposition dimension. However the sizes (i.e., cardinalities) of $\mathfrak{U}^{\text{opt}}(\varepsilon)$ and $\mathfrak{U}^{\text{PW}}(\varepsilon)$ could be very different, especially for $p = \infty$ and/or small a .

For instance, for $p = 2$ we have the following results. In the case of $\varepsilon = 10^{-1}$

$$d^{\text{sup}}(10^{-1}) \leq \begin{cases} 1 & \text{for } a = 4, \\ 1 & \text{for } a = 3, \\ 2 & \text{for } a = 2, \end{cases}$$

and

$$|\mathfrak{U}^{\text{opt}}(10^{-1})| = \begin{cases} 2 & \text{for } a = 4, \\ 2 & \text{for } a = 3, \\ 4 & \text{for } a = 2, \end{cases} \quad \text{whereas} \quad |\mathfrak{U}^{\text{PW}}(10^{-1})| = \begin{cases} 3 & \text{for } a = 4, \\ 5 & \text{for } a = 3, \\ 15 & \text{for } a = 2. \end{cases}$$

For $\varepsilon = 10^{-2}$

$$d^{\text{sup}}(10^{-2}) \leq \begin{cases} 2 & \text{for } a = 4, \\ 2 & \text{for } a = 3, \\ 3 & \text{for } a = 2 \end{cases}$$

and

$$|\mathfrak{U}^{\text{opt}}(10^{-2})| = \begin{cases} 4 & \text{for } a = 4, \\ 7 & \text{for } a = 3, \\ 30 & \text{for } a = 2, \end{cases} \quad \text{whereas} \quad |\mathfrak{U}^{\text{PW}}(10^{-2})| = \begin{cases} 8 & \text{for } a = 4, \\ 18 & \text{for } a = 3, \\ 158 & \text{for } a = 2. \end{cases}$$

And finally, for $\varepsilon = 10^{-3}$

$$d^{\text{sup}}(10^{-3}) \leq \begin{cases} 2 & \text{for } a = 4, \\ 3 & \text{for } a = 3, \\ 4 & \text{for } a = 2, \end{cases}$$

and

$$|\mathfrak{U}^{\text{opt}}(10^{-3})| = \begin{cases} 9 & \text{for } a = 4, \\ 24 & \text{for } a = 3, \\ 255 & \text{for } a = 2, \end{cases} \quad \text{whereas} \quad |\mathfrak{U}^{\text{PW}}(10^{-3})| = \begin{cases} 20 & \text{for } a = 4, \\ 70 & \text{for } a = 3, \\ 1481 & \text{for } a = 2. \end{cases}$$

The results for $a = 4$ suggest that to achieve an error smaller than 10^{-3} it is enough to approximate $f(\mathbf{x}) = \sum_{\mathbf{u} \in \mathbb{N}_+, |\mathbf{u}| < \infty} f_{\mathbf{u}}(\mathbf{x})$ by

$$f_{\emptyset} + f_{\{1\}}(x_1) + \cdots + f_{\{5\}}(x_5) + f_{\{1,2\}}(x_1, x_2) + f_{\{1,3\}}(x_1, x_3) + f_{\{1,4\}}(x_1, x_4).$$

As for the quasi-optimal sets, they are the same for $a = 4$ and slightly larger for $a = 3, 2$:

$$|\mathfrak{U}^{\text{q-opt}}(10^{-1})| = 6 \text{ for } a = 2, \quad |\mathfrak{U}^{\text{q-opt}}(10^{-2})| = 32 \text{ for } a = 2,$$

$$\text{and} \quad |\mathfrak{U}^{\text{q-opt}}(10^{-3})| = \begin{cases} 26 & \text{for } a = 3, \\ 261 & \text{for } a = 2. \end{cases}$$

For $p = \infty$ we have

$$d^{\text{sup}}(10^{-1}) \leq \begin{cases} 1 & \text{for } a = 4, \\ 1 & \text{for } a = 3, \\ 3 & \text{for } a = 2, \end{cases}$$

and

$$|\mathfrak{U}^{\text{opt}}(10^{-1})| = \begin{cases} 2 & \text{for } a = 4, \\ 3 & \text{for } a = 3, \\ 33 & \text{for } a = 2, \end{cases} \quad \text{whereas} \quad |\mathfrak{U}^{\text{PW}}(10^{-1})| = \begin{cases} 7 & \text{for } a = 4, \\ 21 & \text{for } a = 3, \\ 2358 & \text{for } a = 2. \end{cases}$$

Now for $\varepsilon = 10^{-2}$

$$d^{\text{sup}}(10^{-2}) \leq \begin{cases} 2 & \text{for } a = 4, \\ 2 & \text{for } a = 3, \\ 4 & \text{for } a = 2, \end{cases}$$

and

$$|\mathfrak{U}^{\text{opt}}(10^{-2})| = \begin{cases} 5 & \text{for } a = 4, \\ 15 & \text{for } a = 3, \\ 1346 & \text{for } a = 2, \end{cases} \quad \text{whereas} \quad |\mathfrak{U}^{\text{PW}}(10^{-2})| = \begin{cases} 21 & \text{for } a = 4, \\ 149 & \text{for } a = 3, \\ 120,935 & \text{for } a = 2. \end{cases}$$

For $\varepsilon = 10^{-3}$

$$d^{\text{sup}}(10^{-3}) \leq \begin{cases} 2 & \text{for } a = 4, \\ 3 & \text{for } a = 3, \\ 6 & \text{for } a = 2, \end{cases}$$

and

$$|\mathfrak{U}^{\text{opt}}(10^{-3})| = \begin{cases} 15 & \text{for } a = 4, \\ 83 & \text{for } a = 3, \\ 45,446 & \text{for } a = 2, \end{cases} \quad \text{whereas} \quad |\mathfrak{U}^{\text{PW}}(10^{-3})| = \begin{cases} 72 & \text{for } a = 4, \\ 923 & \text{for } a = 3. \end{cases}$$

For the tests above, the quasi-optimal active set was different from the corresponding optimal active set in the following cases only:

$$|\mathfrak{U}^{\text{q-opt}}(10^{-1})| = 38 \text{ for } a = 2, \quad |\mathfrak{U}^{\text{q-opt}}(10^{-2})| = 1904 \text{ for } a = 2,$$

$$\text{and} \quad |\mathfrak{U}^{\text{q-opt}}(10^{-3})| = \begin{cases} 92 & \text{for } a = 3, \\ 52,159 & \text{for } a = 2. \end{cases}$$

A collection of the active sets constructed above have been listed in full in the Appendix.

Our algorithms can also be used to construct the active sets where, instead of the standard worst case error, the normalized worst case error is used. More precisely, for the normalized worst case error we would like to have sets $\mathfrak{U}_{\text{norm}}(\varepsilon)$ such that the integral of $\sum_{\mathbf{u} \notin \mathfrak{U}_{\text{norm}}(\varepsilon)} f_{\mathbf{u}}$ is bounded by

$$\varepsilon \|\mathcal{S}\| \|f\| \quad \text{for all } f \in \mathcal{F}_{\gamma},$$

where $\|\mathcal{S}\|$ is the norm of the integration operator. Since in our case $\|\mathcal{S}\| \geq 1$, the corresponding active sets $\mathfrak{U}_{\text{norm}}^X(\varepsilon)$ are subsets of $\mathfrak{U}^X(\varepsilon)$ for $X \in \{\text{opt}, \text{q-opt}, \text{PW}\}$ and could be even smaller.

2 Basic Definitions

We provide in this section basic concepts and definitions for special spaces of functions that are very often assumed in the literature, especially in the context of quasi-Monte Carlo methods. The presented algorithms can easily be modified to more general spaces.

2.1 γ -Weighted Spaces

We follow here [2]. For $D = [0, 1]$, let $\mathcal{D} = D^{\mathbb{N}_+}$ be the set of sequences (points) $\mathbf{x} = [x_1, x_2, \dots]$ with $x_i \in D$. Here \mathbb{N}_+ is the set of positive integers and we will use \mathbf{u} and \mathbf{v} to denote finite subsets of \mathbb{N}_+ . We will also use the following notation: For $\mathbf{x} \in \mathcal{D}$ and \mathbf{u} , by $[\mathbf{x}_{\mathbf{u}}; \mathbf{0}_{\mathbf{u}^c}]$ we denote the point in \mathcal{D} such that

$$[\mathbf{x}_{\mathbf{u}}; \mathbf{0}_{\mathbf{u}^c}] = [y_1, y_2, \dots] \quad \text{with} \quad y_j = \begin{cases} x_j & \text{if } j \in \mathbf{u}, \\ 0 & \text{if } j \notin \mathbf{u}. \end{cases}$$

Next,

$$f^{(\mathbf{u})} = \prod_{j \in \mathbf{u}} \frac{\partial}{\partial x_j} f.$$

For given

$$p \in [1, \infty],$$

let $\mathcal{F}_{\gamma,p}$ be the Banach space of functions defined on \mathcal{D} with the following norm

$$\|f\|_{\mathcal{F}_{\gamma,p}} = \left(\sum_{\mathbf{u} \in \mathbb{N}_+, |\mathbf{u}| < \infty} \gamma_{\mathbf{u}}^{-p} \|f^{(\mathbf{u})}([\cdot; \mathbf{u}; \mathbf{0}_{\mathbf{u}^c}])\|_{L_p}^p \right)^{1/p}.$$

Of course, for $p = \infty$,

$$\|f\|_{\mathcal{F}_{\gamma,p}} = \sup_{\mathbf{u} \in \mathbb{N}_+, |\mathbf{u}| < \infty} \frac{\|f^{(\mathbf{u})}([\cdot; \mathbf{u}; \mathbf{0}_{\mathbf{u}^c}])\|_{L_\infty}}{\gamma_{\mathbf{u}}}.$$

We assume that the numbers $\gamma_{\mathbf{u}}$ (called *weights*) are of product form (see [8])

$$\gamma_{\mathbf{u}} = \prod_{j \in \mathbf{u}} \frac{c}{j^a} \quad \text{for positive } a \text{ and } c. \quad (1)$$

In general choosing the weights (in our case choosing a and c) for a specific integral or application is a difficult problem which we do not attempt to address here. We assume that the parameters a and c are given with the problem.

It was shown in [2] that any $f \in \mathcal{F}_{\gamma,p}$ admits a unique decomposition, called the *anchored decomposition*,

$$f = \sum_{\mathbf{u} \in \mathbb{N}_+, |\mathbf{u}| < \infty} f_{\mathbf{u}}$$

with $f_{\mathbf{u}}$ given by

$$f_{\mathbf{u}}(\mathbf{x}) = T_{\mathbf{u}}(h_{\mathbf{u}})(\mathbf{x}) := \int_{D^{|\mathbf{u}|}} h_{\mathbf{u}}(\mathbf{t}) \prod_{j \in \mathbf{u}} (x_j - t_j)_+^0 d\mathbf{t} \quad \text{for some } h_{\mathbf{u}} \in L_p(D^{|\mathbf{u}|}),$$

where $(x_j - t_j)_+^0$ is 1 if $x_j > t_j$ and 0 otherwise. The functions $f_{\mathbf{u}}$ belong to the following Banach spaces $F_{\mathbf{u}}$

$$F_{\mathbf{u}} = T_{\mathbf{u}}(L_p) \quad \text{and} \quad \|f_{\mathbf{u}}\|_{F_{\mathbf{u}}} = \|f_{\mathbf{u}}^{(\mathbf{u})}\|_{L_p}.$$

Of course, F_{\emptyset} is the space of constant functions with the absolute value as its norm. The spaces $F_{\mathbf{u}}$ for $\mathbf{u} \neq \emptyset$ are anchored at 0 since $f_{\mathbf{u}}(\mathbf{x}) = 0$ if there is $j \in \mathbf{u}$ with $x_j = 0$. This is why

$$f^{(\mathbf{u})}([\cdot; \mathbf{u}; \mathbf{0}_{\mathbf{u}^c}]) = f_{\mathbf{u}}^{(\mathbf{u})} \quad \text{and} \quad \|f\|_{\mathcal{F}_{\gamma,p}} = \left(\sum_{\mathbf{u} \in \mathbb{N}_+, |\mathbf{u}| < \infty} \gamma_{\mathbf{u}}^{-p} \|f_{\mathbf{u}}\|_{F_{\mathbf{u}}}^p \right)^{1/p}.$$

The space $\mathcal{F}_{\gamma,p}$ contains in particular the following class of functions.

Example 1 For a smooth function $g : \mathbb{R} \rightarrow \mathbb{R}$ and fast decaying numbers a_1, a_2, \dots , consider

$$f(\mathbf{x}) = g\left(\sum_{j=1}^{\infty} x_j a_j\right) \quad \text{for } x_j \in D. \quad (2)$$

Then

$$f^{(\mathbf{u})}([\mathbf{x}_{\mathbf{u}}; \mathbf{0}_{\mathbf{u}^c}]) = g^{(|\mathbf{u}|)} \left(\sum_{j \in \mathbf{u}} x_j a_j \right) \prod_{j \in \mathbf{u}} a_j.$$

Hence $f \in \mathcal{F}_{\gamma, p}$ if the derivatives of g and the coefficients a_j satisfy

$$\left(\sum_{\mathbf{u} \subset \mathbb{N}_+, |\mathbf{u}| < \infty} \frac{\prod_{j \in \mathbf{u}} |a_j|^p}{\gamma_{\mathbf{u}}^p} \int_{D^{|\mathbf{u}|}} \left| g^{(|\mathbf{u}|)} \left(\sum_{j \in \mathbf{u}} x_j a_j \right) \right|^p d\mathbf{x}_{\mathbf{u}} \right)^{1/p} < \infty.$$

2.2 Integration Problem

Consider the following integration functional

$$\mathcal{S} : \mathcal{F}_{\gamma, p} \rightarrow \mathbb{R}$$

given by

$$\mathcal{S}(f) = \lim_{s \rightarrow \infty} \int_{D^s} f(x_1, \dots, x_s, 0, \dots, 0) d[x_1, \dots, x_s].$$

Let p^* denote the conjugate of p ,

$$\frac{1}{p} + \frac{1}{p^*} = 1.$$

We assume that

$$\left(\sum_{\mathbf{u} \subset \mathbb{N}_+, |\mathbf{u}| < \infty} \gamma_{\mathbf{u}}^{p^*} (p^* + 1)^{-|\mathbf{u}|} \right)^{1/p^*} < \infty \quad (3)$$

since the left hand side of (3) is the norm of \mathcal{S} , i.e., (3) is a necessary and sufficient condition for continuity of \mathcal{S} . Indeed, letting $S_{\mathbf{u}}$ be the restriction of \mathcal{S} to $F_{\mathbf{u}}$, we have that

$$\|S_{\mathbf{u}}\|_{F_{\mathbf{u}}} = \sup_{\|f_{\mathbf{u}}\|_{F_{\mathbf{u}}}=1} S_{\mathbf{u}}(f_{\mathbf{u}}) = \frac{1}{(p^* + 1)^{|\mathbf{u}|/p^*}}$$

which, with an application of Hölder's inequality, yields (3). For product weights of the form (1), we have

$$\|\mathcal{S}\| = \left(\sum_{\mathbf{u}} \gamma_{\mathbf{u}}^{p^*} (p^* + 1)^{-|\mathbf{u}|} \right)^{1/p^*} = \prod_{j=1}^{\infty} \left(1 + \frac{c^{p^*}}{j^{ap^*} (p^* + 1)} \right)^{1/p^*}.$$

Hence for product weights, (3) is equivalent to $a > 1/p^*$. For the remainder of the paper it is assumed that $a > 1/p^*$.

A very important part of the *Multivariate Decomposition Method* (MDM for short) is a construction of *active sets* $\mathfrak{U}(\varepsilon)$, i.e., sets that satisfy

$$\left| \mathcal{S} \left(\sum_{\mathbf{u} \notin \mathfrak{U}(\varepsilon)} f_{\mathbf{u}} \right) \right| \leq \varepsilon \left\| \sum_{\mathbf{u} \notin \mathfrak{U}(\varepsilon)} f_{\mathbf{u}} \right\|_{\mathcal{F}_{\gamma, p}} \quad \text{for all } f \in \mathcal{F}_{\gamma, p}. \quad (4)$$

The essence of (4) is that, when approximating $\mathcal{S}(f)$, it is enough to restrict the attention to functions

$$\sum_{\mathbf{u} \in \mathfrak{U}(\varepsilon)} f_{\mathbf{u}},$$

since any algorithm approximating $\sum_{\mathbf{u} \in \mathfrak{U}(\varepsilon)} S_{\mathbf{u}}(f_{\mathbf{u}})$ with the worst case error on $\bigoplus_{\mathbf{u} \in \mathfrak{U}(\varepsilon)} F_{\mathbf{u}}$ bounded by ε has its worst case error on the whole space $\mathcal{F}_{\gamma,p}$ bounded by

$$2^{1/p^*} \varepsilon.$$

The factor of $2^{1/p^*}$ is the result of applying Hölder's inequality, see, e.g., [3]. Clearly, there are many sets satisfying (4), and we would like to construct possibly small active sets.

Definition 2 *We say that an active set, denoted by $\mathfrak{U}^{\text{opt}}(\varepsilon)$ is optimal, if*

$$|\mathfrak{U}^{\text{opt}}(\varepsilon)| = \min\{|\mathfrak{U}(\varepsilon)| : \mathfrak{U}(\varepsilon) \text{ satisfies (4)}\}.$$

We also define the ε -superposition dimension as the smallest $d(\mathfrak{U}(\varepsilon))$ among all active sets,

$$d^{\text{sup}}(\varepsilon) := \min\{d(\mathfrak{U}(\varepsilon)) : \mathfrak{U}(\varepsilon) \text{ satisfies (4)}\}.$$

3 Constructing Active Sets $\mathfrak{U}(\varepsilon)$

A construction of active sets was first proposed in [7]. The corresponding sets will be denoted by $\mathfrak{U}^{\text{PW}}(\varepsilon)$. It was shown there that

$$d(\mathfrak{U}^{\text{PW}}(\varepsilon)) = O\left(\frac{\ln(1/\varepsilon)}{\ln(\ln(1/\varepsilon))}\right) \quad \text{as } \varepsilon \rightarrow 0$$

and that the cardinality of $\mathfrak{U}^{\text{PW}}(\varepsilon)$ is polynomial in $1/\varepsilon$. It is easy to see that for $p = 1$, $\mathfrak{U}^{\text{PW}}(\varepsilon)$ are optimal. However, as we shall see, their size might be too large for large values of p , especially for $p = \infty$.

Let \mathfrak{U} be a set of subsets \mathbf{u} . Then

$$\begin{aligned} \left| \mathcal{S}\left(\sum_{\mathbf{u} \notin \mathfrak{U}} f_{\mathbf{u}}\right) \right| &\leq \sum_{\mathbf{u} \notin \mathfrak{U}} \|f_{\mathbf{u}}\|_{F_{\mathbf{u}}} \|S_{\mathbf{u}}\|_{F_{\mathbf{u}}} = \sum_{\mathbf{u} \notin \mathfrak{U}} \frac{\|f_{\mathbf{u}}\|_{F_{\mathbf{u}}}}{\gamma_{\mathbf{u}}} \gamma_{\mathbf{u}} \|S_{\mathbf{u}}\|_{F_{\mathbf{u}}} \\ &\leq \left\| \sum_{\mathbf{u} \notin \mathfrak{U}} f_{\mathbf{u}} \right\|_{\mathcal{F}_{\gamma,p}} \left(\sum_{\mathbf{u} \notin \mathfrak{U}} \gamma_{\mathbf{u}}^{p^*} \|S_{\mathbf{u}}\|_{F_{\mathbf{u}}}^{p^*} \right)^{1/p^*}. \end{aligned}$$

Hence we are looking for $\mathfrak{U}(\varepsilon)$ such that

$$\left(\sum_{\mathbf{u} \notin \mathfrak{U}(\varepsilon)} \gamma_{\mathbf{u}}^{p^*} \|S_{\mathbf{u}}\|_{F_{\mathbf{u}}}^{p^*} \right)^{1/p^*} \leq \varepsilon. \quad (5)$$

Since Hölder's inequality is sharp, (5) is equivalent to (4).

For the sake of completeness, we recall the construction for $p = 1$, see [7].

3.1 Case $p = 1$

For $p = 1$, we have $p^* = \infty$ and $\|S_{\mathbf{u}}\|_{F_{\mathbf{u}}} = 1$ for all \mathbf{u} . Hence

$$\left(\sum_{\mathbf{u} \notin \mathfrak{U}(\varepsilon)} \gamma_{\mathbf{u}}^{p^*} \|S_{\mathbf{u}}\|_{F_{\mathbf{u}}}^{p^*} \right)^{1/p^*} = \sup_{\mathbf{u} \notin \mathfrak{U}(\varepsilon)} \gamma_{\mathbf{u}}$$

which for product weights reduces to $\sup_{\mathbf{u} \notin \mathfrak{U}(\varepsilon)} \prod_{j \in \mathbf{u}} c j^{-a}$. Therefore

$$\mathfrak{U}^{\text{PW}}(\varepsilon) = \left\{ \mathbf{u} : \prod_{j \in \mathbf{u}} \frac{c}{j^a} > \varepsilon \right\}. \quad (6)$$

It is easy to see that $\mathfrak{U}^{\text{PW}}(\varepsilon)$ is the smallest set satisfying (4), i.e., it is a subset of any $\mathfrak{U}(\varepsilon)$ satisfying (5).

The examples of \mathfrak{U}^{PW} for specific values of a and ε are presented in the Appendix. For simplicity we use $c = 1$ there.

3.2 Case of $p > 1$

For $p > 1$, the conjugate p^* is finite and the construction of $\mathfrak{U}(\varepsilon)$ is more complicated.

We begin by recalling the construction of $\mathfrak{U}^{\text{PW}}(\varepsilon)$ in [7]. To simplify the notation, let

$$\bar{\gamma}_{\mathbf{u}} = \frac{\gamma_{\mathbf{u}}^{p^*}}{(p^* + 1)^{|\mathbf{u}|}} = \left(\frac{c^{p^*}}{p^* + 1} \right)^{|\mathbf{u}|} \prod_{j \in \mathbf{u}} j^{-a p^*}.$$

For given ε and p , a special threshold is computed and all \mathbf{u} with $\bar{\gamma}_{\mathbf{u}}$ exceeding the threshold are included in the active set. More precisely, for $t \in (1/(ap^*), 1)$ a threshold is given by

$$\text{Threshold}(\varepsilon, t) = \left(\frac{\varepsilon^{p^*}}{\sum_{\mathbf{u} \subset \mathbb{N}_+, |\mathbf{u}| < \infty} \bar{\gamma}_{\mathbf{u}}^t} \right)^{1/(1-t)}.$$

Note that the interval $(1/(ap^*), 1)$ is non-empty by the assumption that $a > 1/p^*$ introduced in Section 2.2. In our numerical experiments we approximated the sum of $\bar{\gamma}_{\mathbf{u}}^t$ for $t = i/40$ ($39 \geq i > 40/(ap^*)$) and selected the value which resulted in the largest $\text{Threshold}(\varepsilon, t)$. The approximations are calculated in a similar way as the computation of A_s explained later (see (7)).

Clearly, $\mathfrak{U}^{\text{PW}}(\varepsilon)$ contains a number of \mathbf{u} 's with the largest $\bar{\gamma}_{\mathbf{u}}$; however, the number of them could be much larger than needed.

The collection of those \mathbf{u} with the largest $\bar{\gamma}_{\mathbf{u}}$ that are necessary for (4) would result in the optimal set $\mathfrak{U}^{\text{opt}}(\varepsilon)$. Since this optimal set is always a subset of $\mathfrak{U}^{\text{PW}}(\varepsilon)$, the good property

$$d(\mathfrak{U}^{\text{opt}}(\varepsilon)) = O\left(\frac{\ln(1/\varepsilon)}{\ln(\ln(1/\varepsilon))}\right) \quad \text{as } \varepsilon \rightarrow 0,$$

is preserved. More precisely, let $(\mathbf{u}_j)_{j \in \mathbb{N}_+}$ be a sequence of all subsets \mathbf{u} ordered so that

$$\bar{\gamma}_{\mathbf{u}_j} \geq \bar{\gamma}_{\mathbf{u}_{j+1}} \quad j = 1, 2, \dots$$

Then

$$\mathfrak{U}^{\text{opt}}(\varepsilon) = \{\mathbf{u}_1, \dots, \mathbf{u}_k\}$$

with $k = k(\varepsilon)$ such that

$$\|\mathcal{S}\|^{p^*} - \sum_{j=1}^k \bar{\gamma}_{\mathbf{u}_j} \leq \varepsilon^{p^*} < \|\mathcal{S}\|^{p^*} - \sum_{j=1}^{k-1} \bar{\gamma}_{\mathbf{u}_j}.$$

The problem with this approach is that we do not know *a priori* the number $k = k(\varepsilon)$ and ordering a large number of $\bar{\gamma}_{\mathbf{u}}$ might be too expensive. This is why the numbers $\bar{\gamma}_{\mathbf{u}}$ will be ordered on-line. Actually, we propose two ways of constructing active sets. The first and simpler one produces what we call, *quasi-optimal* sets $\mathfrak{U}^{\text{q-opt}}(\varepsilon)$ and it uses a partial ordering of $\bar{\gamma}_{\mathbf{u}}$. The second one, uses ordering of $\bar{\gamma}_{\mathbf{u}}$ and produces optimal sets $\mathfrak{U}^{\text{opt}}(\varepsilon)$. However, as we will see the difference between both sets is very small; sometimes these sets are equal.

The numbers $\bar{\gamma}_{\mathbf{u}}$ have the following properties that are crucial for our construction of quasi-optimal and optimal sets. Let ℓ be a given cardinality. In what follows we will write

$$\mathbf{u} = \{u_1, \dots, u_\ell\}, \quad \text{where } u_1 < \dots < u_\ell.$$

The first property is: If

$$\mathbf{u} = \{u_1, \dots, u_\ell\} \text{ and } \mathbf{v} = \{v_1, \dots, v_\ell\} \text{ with } v_j \geq u_j \text{ for all } j$$

then

$$\bar{\gamma}_{\mathbf{u}} \geq \bar{\gamma}_{\mathbf{v}}.$$

The other property is: For $\ell + 1 \geq c^{1/a}$,

$$\bar{\gamma}_{\{u_1, \dots, u_\ell\}} \geq \bar{\gamma}_{\{u_1, \dots, u_\ell, u_{\ell+1}\}}.$$

We are ready to describe the constructions of active sets. First we need to approximate

$$A = \sum_{\mathbf{u} \in \mathbb{N}_+, |\mathbf{u}| < \infty} \bar{\gamma}_{\mathbf{u}}$$

from above and with the relative error significantly smaller than ε^{p^*} . This can be done as follows. For a large natural number s

$$\begin{aligned} A &= \exp \left(\ln \left(\prod_{j=s+1}^{\infty} \left(1 + \frac{(c/j^a)^{p^*}}{p^* + 1} \right) \right) \right) \prod_{j=1}^s \left(1 + \frac{(c/j^a)^{p^*}}{p^* + 1} \right) \\ &\leq \exp \left(\frac{c^{p^*}}{p^* + 1} \sum_{j=s+1}^{\infty} j^{-a p^*} \right) \prod_{j=1}^s \left(1 + \frac{(c/j^a)^{p^*}}{p^* + 1} \right) \\ &\leq \exp \left(\frac{c^{p^*}}{p^* + 1} \int_{s+1/2}^{\infty} x^{-a p^*} dx \right) \prod_{j=1}^s \left(1 + \frac{(c/j^a)^{p^*}}{p^* + 1} \right) \end{aligned}$$

$$= \exp \left(\frac{c^{p^*}}{(p^* + 1)(ap^* - 1)(s + 1/2)^{ap^* - 1}} \right) \prod_{j=1}^s \left(1 + \frac{(c/j^a)^{p^*}}{p^* + 1} \right) =: A_s. \quad (7)$$

It is easy to see that the relative error between A and its approximation A_s is proportional to $1/s^{2a p^* - 2}$ with the asymptotic constant $c^{p^*} / ((p^* + 1) 2^{ap^* - 1} \prod_{j=1}^{\infty} (1 + (c/j^a)^{p^*} / (p^* + 1)))$.

A general idea of our construction is to select sets \mathbf{u} with large $\bar{\gamma}_{\mathbf{u}}$ and subtract $\bar{\gamma}_{\mathbf{u}}$ from A_s . This is repeated until A_s is reduced to or below ε^{p^*} .

More specifically, consider a partition of \mathbb{R}_+ into intervals I_i such that the numbers in I_j are greater than those in I_{j+1} . For simplicity, we used

$$I_1 = [10^{-1}, \infty), \quad \text{and} \quad I_j = [10^{-j}, 10^{-j+1}) \quad \text{for } j = 2, 3, \dots$$

in our numerical experiments when constructing quasi-optimal sets. However, we think that a better partition is possible, especially when constructing optimal active sets. We also associate with every interval a list L_j that contains those \mathbf{u} for which $\bar{\gamma}_{\mathbf{u}}$ has been subtracted from A_s in j th step.

In the first $j = 1$ step, add the empty set to L_1 and subtract $\bar{\gamma}_{\emptyset} = 1$ from A_s . If the new A_s satisfies $A_s \leq \varepsilon^{p^*}$, then terminate. Otherwise consider non-empty sets \mathbf{u} in the order of increasing cardinalities. Hence start with singleton sets $\mathbf{u} = \{i\}$ for $i = 1, \dots, k$, where k is the largest integer such that $\bar{\gamma}_{\{k\}}$ is in I_1 . Place $\{k + 1\}$ into list L_2 , and start subtracting from A_s the values $\bar{\gamma}_{\{i\}}$ and store $\{i\}$ in L_1 until either the difference becomes less than or equal to ε^{p^*} , in which case we terminate, or $i = k$. Next repeat the same for sets of cardinality 2, starting with sets $\{1, i\}$ for $i \leq k$, where now k is the largest integer such that $\bar{\gamma}_{\{1, k\}} \in I_1$. Store $\{1, k + 1\}$ in L_2 . Next consider sets $\{2, i\}$, $\{3, i\}$, etc., until either all cardinality 2 sets corresponding to the current interval have been visited or the new value of A_s is $\leq \varepsilon^{p^*}$, in which case we terminate. Continue working through the sets in order of increasing cardinality ℓ until $\bar{\gamma}_{\{1, 2, \dots, \ell\}} \notin I_1$ and $\ell \geq c$ (of course, ℓ is always at least c for $c \leq 1$). Then move to step $j = 2$. The procedure in this (and later) steps is very similar except that for a fixed cardinality of \mathbf{u} , we check if any such set has already been placed in L_2 in the 1st step. If it has, we start working with such sets first. For instance, for cardinality 1, if $\{k + 1\} \in L_2$, then we begin with sets $\{i\}$ for $i \geq k + 1$. For cardinality 2, if $\{i_1, i_2\}$ has been placed in L_2 , then we inspect sets $\{i_1, i\}$ for $i \geq i_2$, before any other sets of cardinality two are considered. Once we find $\bar{\gamma}_{\{i, i+1\}} \notin I_2$, we store $\{i, i + 1\}$ in L_3 and proceed to sets of cardinality 3, etc.

At the very end, $\mathfrak{U}^{\text{q-opt}}(\varepsilon)$ consists of all subsets \mathbf{u} whose values $\bar{\gamma}_{\mathbf{u}}$ were subtracted from A_s . The main procedure is outlined in Algorithm 2. In all of the algorithms j_{\max} and ℓ_{\max} are computational thresholds denoting, respectively, the maximum number of intervals to be searched through and the maximum allowed cardinality of sets.

To search through the sets in a systematic way, we must keep track of the current set, \mathbf{u} , and the index, i , that we are incrementing from. The subroutine **increment-u** outlined below in Algorithm 1 details how to increment \mathbf{u} from index i .

Algorithm 1 (Subroutine: **increment-u**)

inputs: \mathbf{u}, i **output:** \mathbf{u}

- 1: $u_i \leftarrow u_i + 1$ ▷ updating u_i first
 - 2: **for** $r = i + 1, i + 2, \dots, |\mathbf{u}|$ **do** ▷ incrementing \mathbf{u} from index $i + 1$
 - 3: $u_r \leftarrow u_i + r - i$
 - 4: **end for**
 - 5: **return** \mathbf{u}
-

Algorithm 2 (Constructing the quasi-optimal active set)

inputs: $\varepsilon, p^*, s, (\bar{\gamma}_{\mathbf{u}})_{|\mathbf{u}| < \infty}, (I_j)_{j=1}^{j_{\max}}$ **output:** $\mathcal{U}^{\text{q-opt}}(\varepsilon)$

- 1: $L_j \leftarrow \emptyset$ for all $j = 1, 2, \dots, j_{\max}$ ▷ initialising
 - 2: $\mathcal{U}^{\text{q-opt}}(\varepsilon) \leftarrow \{\emptyset\}$
 - 3: $T \leftarrow A_s - \varepsilon^{p^*} - \bar{\gamma}_{\emptyset}$ ▷ T tracks difference between $A_s - \varepsilon^{p^*}$ and weights
 - 4: **if** $T \leq 0$ **then** **return** $\mathcal{U}^{\text{q-opt}}(\varepsilon)$ ▷ quasi-optimal active set is complete
 - 5: **for** $j = 1, 2, \dots, j_{\max}$ **do** ▷ looping over intervals
 - 6: ▷ first handle sets found at previous step
 - 7: $(\mathcal{U}^{\text{q-opt}}(\varepsilon), T, \ell_{\text{next}}, L_{j+1}) \leftarrow \mathbf{q-opt-search}(\mathcal{U}^{\text{q-opt}}(\varepsilon), T, (\bar{\gamma}_{\mathbf{u}})_{|\mathbf{u}| < \infty}, L_j, L_{j+1}, I_j)$
 - 8: **if** $T \leq 0$ **then** **return** $\mathcal{U}^{\text{q-opt}}(\varepsilon)$ ▷ quasi-optimal active set is complete
 - 9: **for** $\ell = \ell_{\text{next}}, \ell_{\text{next}} + 1, \dots, \ell_{\max}$ **do** ▷ search through unvisited sets
 - 10: $\mathbf{u} = \{1, 2, \dots, \ell\}$
 - 11: $i \leftarrow \ell$ ▷ i keeps track of index to increment \mathbf{u} from
 - 12: **if** $\bar{\gamma}_{\mathbf{u}} \notin I_j$ **and** $\ell \geq c$ **then** **break** ▷ no more \mathbf{u} with $\bar{\gamma}_{\mathbf{u}} \in I_j$
 - 13: **while** $i > 0$ **do** ▷ when $i = 0$ there are no more \mathbf{u} of cardinality ℓ
 - 14: **if** $\bar{\gamma}_{\mathbf{u}} \in I_j$ **then**
 - 15: add \mathbf{u} to $\mathcal{U}^{\text{q-opt}}(\varepsilon)$
 - 16: $T \leftarrow T - \bar{\gamma}_{\mathbf{u}}$
 - 17: **if** $T \leq 0$ **then** **return** $\mathcal{U}^{\text{q-opt}}(\varepsilon)$ ▷ quasi-optimal set is complete
 - 18: $i \leftarrow \ell$ ▷ continue incrementing from last index
 - 19: **else**
 - 20: add \mathbf{u} to L_{j+1}
 - 21: $i \leftarrow i - 1$ ▷ start incrementing from lower index
 - 22: **if** $i = 0$ **then** **break** ▷ go to next cardinality
 - 23: **end if**
 - 24: $\mathbf{u} \leftarrow \mathbf{increment-u}(\mathbf{u}, i)$
 - 25: **end while**
 - 26: **end for**
 - 27: **end for**
-

To make the presentation clearer Algorithm 2 is broken into two parts: First, we search starting from the sets found in the previous interval, which is handled by the subroutine **q-opt-search** in Algorithm 3. Then we continue searching through sets in order of increasing cardinality (line 9) starting where **q-opt-search** finished, at cardinality ℓ_{next} .

The basic search structure is the same, however in **q-opt-search** each set we visit is checked to reduce multiple visits to a single set and ensure that the same set is not added to $\mathfrak{U}^{\text{q-opt}}(\varepsilon)$ more than once.

The notation

$$(\mathfrak{U}^{\text{q-opt}}(\varepsilon), T, \ell_{\text{next}}, L_{j+1}) \leftarrow \mathbf{q-opt-search}(\mathfrak{U}^{\text{q-opt}}(\varepsilon), T, (\bar{\gamma}_{\mathbf{u}})_{|\mathbf{u}| < \infty}, L_j, L_{j+1}, I_j),$$

denotes that we call **q-opt-search** with inputs $\mathfrak{U}^{\text{q-opt}}(\varepsilon)$, T , $(\bar{\gamma}_{\mathbf{u}})_{|\mathbf{u}| < \infty}$, L_j , L_{j+1} , I_j and then use the output to update $\mathfrak{U}^{\text{q-opt}}(\varepsilon)$, T , ℓ_{next} and L_{j+1} .

Algorithm 3 (Subroutine: **q-opt-search**)

inputs: $\mathfrak{U}^{\text{q-opt}}(\varepsilon)$, T , $(\bar{\gamma}_{\mathbf{u}})_{|\mathbf{u}| < \infty}$, L_j , L_{j+1} , I_j

outputs: $\mathfrak{U}^{\text{q-opt}}(\varepsilon)$, ℓ_{next} , T , L_{j+1}

```

1:  $\ell_{\text{next}} = 1$ 
2: for  $\mathbf{u} \in L_j$  do
3:    $i \leftarrow |\mathbf{u}|$ 
4:   while  $i > 0$  do
5:      $L_j \leftarrow L_j \setminus \mathbf{u}$  ▷ reducing the double-handling of sets
6:     if  $\bar{\gamma}_{\mathbf{u}} \in I_j$  then
7:       if  $\mathbf{u} \in \mathfrak{U}^{\text{q-opt}}(\varepsilon)$  then break ▷ already visited  $\mathbf{u}$  and any future increments
8:       add  $\mathbf{u}$  to  $\mathfrak{U}^{\text{q-opt}}(\varepsilon)$ 
9:        $T \leftarrow T - \bar{\gamma}_{\mathbf{u}}$ 
10:      if  $T \leq 0$  then return  $(\mathfrak{U}^{\text{q-opt}}(\varepsilon), \ell_{\text{next}}, T, L_{j+1})$ 
11:       $i = |\mathbf{u}|$ 
12:    else
13:      add  $\mathbf{u}$  to  $L_{j+1}$  ▷  $\mathbf{u}$  to be checked first in next interval
14:       $i \leftarrow i - 1$ 
15:      if  $i = 0$  then break ▷ go to next  $\mathbf{u} \in L_j$ 
16:    end if
17:     $\mathbf{u} \leftarrow \text{increment-}\mathbf{u}(\mathbf{u}, i)$ 
18:  end while
19:   $\ell_{\text{next}} = |\mathbf{u}| + 1$  ▷ main search will start at cardinality  $|\mathbf{u}| + 1$ 
20: end for
21: return  $(\mathfrak{U}^{\text{q-opt}}(\varepsilon), \ell_{\text{next}}, T, L_{j+1})$ 

```

The construction of optimal active sets is very similar. The main difference is that in the j th step, we first create the list L_j^{unsorted} , order its elements $\mathbf{u} \in L_j^{\text{unsorted}}$ according to decreasing values of $\bar{\gamma}_{\mathbf{u}}$, and next start subtracting the values $\bar{\gamma}_{\mathbf{u}}$ from A_s . Again the lists L_j will hold the sets visited in the previous interval.

In fact, if we do not care whether or not all of the sets are ordered but only that $\mathfrak{U}^{\text{opt}}(\varepsilon)$ consists of the sets with the largest weights, then we only need to sort the sets which come from the final interval. This is because at the previous intervals all of the sets will need to be added to $\mathfrak{U}^{\text{opt}}(\varepsilon)$, regardless of sorting. To do this in practice, for each interval we store the sum of all the weights corresponding to that interval. In Algorithm 4 we denote this by T_j . At the end of the j th step, we check whether I_j is the final interval,

i.e., if $A_s - \sum_{i=1}^j T_i \leq \varepsilon^{p^*}$, if so we sort the sets and add them one-by-one until the active set is complete. Otherwise we add all of the sets in L_j^{unsorted} to $\mathcal{U}^{\text{opt}}(\varepsilon)$ and go to the next interval. For completeness, the construction of optimal active sets is detailed separately below in Algorithm 4 and the subroutine **opt-search** in Algorithm 5.

Algorithm 4 (Constructing the optimal active set)

inputs: $\varepsilon, p^*, s, (\bar{\gamma}_u)_{|u|<\infty}, (I_j)_{j=1}^{j_{\max}}$

output: $\mathcal{U}^{\text{opt}}(\varepsilon)$

```

1:  $T_j \leftarrow 0, L_j^{\text{unsorted}} \leftarrow \emptyset$  and  $L_j \leftarrow \emptyset$  for all  $j = 1, 2, \dots, j_{\max}$  ▷ initialising
2:  $\mathcal{U}^{\text{opt}}(\varepsilon) \leftarrow \{\emptyset\}$ 
3:  $T \leftarrow A_s - \varepsilon^{p^*} - \bar{\gamma}_{\emptyset}$  ▷  $T$  tracks difference between  $A_s - \varepsilon^{p^*}$  and weights
4: if  $T \leq 0$  then return  $\mathcal{U}^{\text{opt}}(\varepsilon)$  ▷ optimal active set is complete
5: for  $j = 1, 2, \dots, j_{\max}$  do ▷ looping over intervals
6: ▷ first handle sets found at previous step
7:  $(\ell_{\text{next}}, T_j, L_j^{\text{unsorted}}, L_{j+1}) \leftarrow \text{opt-search}(\mathcal{U}^{\text{opt}}(\varepsilon), T_j, (\bar{\gamma}_u)_{|u|<\infty}, L_j^{\text{unsorted}}, L_{j+1}, I_j)$ 
8: for  $\ell = \ell_{\text{next}}, \ell_{\text{next}} + 1, \dots, \ell_{\max}$  do ▷ search through unvisited sets
9:    $\mathbf{u} = \{1, 2, \dots, \ell\}$ 
10:    $i \leftarrow \ell$  ▷  $i$  keeps track of index to increment  $\mathbf{u}$  from
11:   if  $\bar{\gamma}_{\mathbf{u}} \notin I_j$  and  $\ell \geq c$  then break ▷ no more  $\mathbf{u}$  with  $\bar{\gamma}_{\mathbf{u}} \in I_j$ 
12:   while  $i > 0$  do ▷ when  $i = 0$  there are no more  $\mathbf{u}$  of cardinality  $\ell$ 
13:     if  $\bar{\gamma}_{\mathbf{u}} \in I_j$  then
14:       add  $\mathbf{u}$  to  $L_j^{\text{unsorted}}$ 
15:        $T_j \leftarrow T_j + \bar{\gamma}_{\mathbf{u}}$ 
16:        $i \leftarrow \ell$  ▷ continue incrementing from last index
17:     else
18:       add  $\mathbf{u}$  to  $L_{j+1}$ 
19:        $i \leftarrow i - 1$  ▷ start incrementing from lower index
20:       if  $i = 0$  then break ▷ go to next cardinality
21:     end if
22:      $\mathbf{u} \leftarrow \text{increment-u}(\mathbf{u}, i)$ 
23:   end while
24: end for
25: if  $T_j \geq T$  then ▷ sorting step, first check if  $I_j$  is the last interval
26:   sort  $L_j^{\text{unsorted}}$ 
27:   for  $\mathbf{u} \in L_j^{\text{sorted}}$  do ▷ add sorted sets until active set is complete
28:     add  $\mathbf{u}$  to  $\mathcal{U}^{\text{opt}}(\varepsilon)$ 
29:      $T \leftarrow T - \bar{\gamma}_{\mathbf{u}}$ 
30:     if  $T \leq 0$  then return  $\mathcal{U}^{\text{opt}}(\varepsilon)$  ▷ optimal active set is complete
31:   end for
32: else ▷ add all sets for the current interval and continue search
33:   add all  $\mathbf{u}$  to  $\mathcal{U}^{\text{opt}}(\varepsilon)$ 
34:    $T \leftarrow T - T_j$ 
35: end if
36: end for

```

Algorithm 5 (Subroutine: **opt-search**)

inputs: $\mathcal{U}^{\text{opt}}(\varepsilon)$, T_j , $(\bar{\gamma}_{\mathbf{u}})_{|\mathbf{u}| < \infty}$, L_j^{unsorted} , L_{j+1} , I_j **outputs:** ℓ_{next} , T_j , L_j^{unsorted} , L_{j+1}

```
1:  $\ell_{\text{next}} = 1$ 
2: for  $\mathbf{u} \in L_j$  do
3:    $i \leftarrow |\mathbf{u}|$ 
4:   while  $i > 0$  do
5:      $L_j \leftarrow L_j \setminus \mathbf{u}$  ▷ reducing the double-handling of sets
6:     if  $\bar{\gamma}_{\mathbf{u}} \in I_j$  then
7:       if  $\mathbf{u} \in L_j^{\text{unsorted}}$  then break ▷ already visited  $\mathbf{u}$  and any future increments
8:       add  $\mathbf{u}$  to  $L_j^{\text{unsorted}}$ 
9:        $T_j \leftarrow T_j + \bar{\gamma}_{\mathbf{u}}$ 
10:       $i = |\mathbf{u}|$ 
11:    else
12:      add  $\mathbf{u}$  to  $L_{j+1}$  ▷  $\mathbf{u}$  to be checked first in next interval
13:       $i \leftarrow i - 1$ 
14:      if  $i = 0$  then break ▷ go to next  $\mathbf{u} \in L_j$ 
15:    end if
16:     $\mathbf{u} \leftarrow \text{increment-}\mathbf{u}(\mathbf{u}, i)$ 
17:  end while
18:   $\ell_{\text{next}} = |\mathbf{u}| + 1$  ▷ main search will start at cardinality  $|\mathbf{u}| + 1$ 
19: end for
20: return  $(\ell_{\text{next}}, T_j, L_j^{\text{unsorted}}, L_{j+1})$ 
```

4 Discussion

In this paper we have introduced the notion of *superposition dimension* and *optimal* active sets to be used in the MDM for multivariate integration and presented an algorithm detailing their construction. We also introduced a second simplified, computationally less intensive version of the algorithm, which constructs *quasi-optimal* active sets. Our numerical results show that the quasi-optimal active sets are of a similar size to the optimal active sets. Often the two sets are exactly the same. In all of our numerical results the optimal and quasi-optimal active sets are smaller than, and have superposition dimension less than or equal to, the active sets using the construction in [7].

To observe how different choices of parameters a and c affect our construction, statistics on the resulting optimal active sets are given in Tables 1-4. Tables 1 and 2 give, respectively, the size and the superposition dimension of the optimal active set for $p = 2$ and an error request of 10^{-2} . For $p = \infty$, $\varepsilon = 10^{-2}$ the size and superposition dimension of the optimal active sets are given in Tables 3 and 4. The results for the quasi-optimal active set are again very similar and so have not been included here. As expected these results demonstrate that as the decay of the weights is slower or the weights become larger (a smaller and c larger) the problem becomes more difficult and the active sets are by necessity larger. However the superposition dimension remains relatively small, at most 6.

	a		
c	4	3	2
$\frac{1}{2}$	3	5	12
1	4	7	30
2	6	14	122

Table 1: $|\mathfrak{U}^{\text{opt}}(10^{-2})|$ for $p = 2$ and different a, c .

	a		
c	4	3	2
$\frac{1}{2}$	1	2	2
1	2	2	3
2	2	3	4

Table 2: $d(\mathfrak{U}^{\text{opt}}(10^{-2}))$ for $p = 2$ and different a, c .

	a		
c	4	3	2
$\frac{1}{2}$	4	7	150
1	5	15	1346
2	8	43	31,013

Table 3: $|\mathfrak{U}^{\text{opt}}(10^{-2})|$ for $p = \infty$ and different a, c .

	a		
c	4	3	2
$\frac{1}{2}$	2	3	4
1	2	2	4
2	2	3	6

Table 4: $d(\mathfrak{U}^{\text{opt}}(10^{-2}))$ for $p = \infty$ and different a, c .

Finally, we have constructed the sets $\mathfrak{U}^{\text{q-opt}}(\varepsilon)$, $\mathfrak{U}^{\text{opt}}(\varepsilon)$, and $\mathfrak{U}^{\text{PW}}(\varepsilon)$ for the product weights with $c = 1$. They are listed explicitly in the Appendix.

Acknowledgements

The authors would like to thank the two anonymous referees for their constructive comments which helped improve the paper.

Appendix

We list here a selection of the constructed active sets from the previous sections (the very largest sets have been omitted). To save the space sometimes we write $[\dots\{x_1, \dots, x_k, x_{k+1}\}]$ to denote the sequence of sets,

$$[\dots\{x_1, \dots, x_k, x_{k+1}\}] = \{x_1, \dots, x_k, x_k + 1\}, \dots, \{x_1, \dots, x_k, x_{k+1}\}.$$

For instance $[\dots\{3\}]$ denotes $\{1\}, \{2\}, \{3\}$ and $[\dots\{1, 5\}]$ denotes $\{1, 2\}, \{1, 3\}, \{1, 4\}, \{1, 5\}$.

Case $p = 1$ and $a = 4$

$$\mathfrak{U}^{\text{PW}}(10^{-1}) = \{\emptyset, \{1\}\}, \quad \mathfrak{U}^{\text{PW}}(10^{-2}) = \{\emptyset, [\dots\{3\}], \{1, 2\}, \{1, 3\}\}, \quad \mathfrak{U}^{\text{PW}}(10^{-3}) = \{\emptyset, [\dots\{5\}], [\dots\{1, 5\}]\}.$$

Case of $p = 1$ and $a = 3$

$$\mathfrak{U}^{\text{PW}}(10^{-1}) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}, \quad \mathfrak{U}^{\text{PW}}(10^{-2}) = \{\emptyset, [\dots\{4\}], [\dots\{1, 4\}]\},$$

$$\mathfrak{U}^{\text{PW}}(10^{-3}) = \{\emptyset, [\dots\{9\}], [\dots\{1, 9\}], \{2, 3\}, \{2, 4\}, \{1, 2, 3\}, \{1, 2, 4\}\}.$$

Case $p = 1$ and $a = 2$

$$\mathfrak{U}^{\text{PW}}(10^{-1}) = \{\emptyset, [\dots\{3\}], \{1, 2\}, \{1, 3\}\},$$

$$\mathfrak{U}^{\text{PW}}(10^{-2}) = \{\emptyset, [\dots\{9\}], [\dots\{1, 9\}], \{2, 3\}, \{2, 4\}, \{1, 2, 3\}, \{1, 2, 4\},$$

$$\begin{aligned} \mathfrak{U}^{\text{PW}}(10^{-3}) = & \{\emptyset, [\dots\{31\}], [\dots\{1, 31\}], [\dots\{2, 15\}], [\dots\{3, 10\}], [\dots\{4, 7\}], \{5, 6\}, [\dots\{1, 2, 15\}], \\ & [\dots\{1, 3, 10\}], [\dots\{1, 4, 7\}], \{1, 5, 6\}, \{2, 3, 4\}, \{2, 3, 5\}, \{1, 2, 3, 4\}, \{1, 2, 3, 5\}\}. \end{aligned}$$

Case $p = 2$ and $a = 4$

$$\mathfrak{U}^{\text{q-opt}}(10^{-1}) = \mathfrak{U}^{\text{opt}}(10^{-1}) = \{\emptyset, \{1\}\}, \quad \text{and} \quad \mathfrak{U}^{\text{PW}}(10^{-1}) = \{\emptyset, \{1\}, \{2\}\},$$

$$\mathfrak{U}^{\text{q-opt}}(10^{-2}) = \mathfrak{U}^{\text{opt}}(10^{-2}) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\} \quad \text{and} \quad \mathfrak{U}^{\text{PW}}(10^{-2}) = \{\emptyset, [\dots\{4\}], [\dots\{1, 4\}]\},$$

$$\mathfrak{U}^{\text{q-opt}}(10^{-3}) = \mathfrak{U}^{\text{opt}}(10^{-3}) = \{\emptyset, [\dots\{5\}], [\dots\{1, 4\}]\}$$

$$\mathfrak{U}^{\text{PW}}(10^{-3}) = \{\emptyset, [\dots\{9\}], [\dots\{1, 8\}], \{2, 3\}, \{2, 4\}, \{1, 2, 3\}\}.$$

Case of $p = 2$ and $a = 3$

$$\mathfrak{U}^{\text{q-opt}}(10^{-1}) = \mathfrak{U}^{\text{opt}}(10^{-1}) = \{\emptyset, \{1\}\} \quad \text{and} \quad \mathfrak{U}^{\text{PW}}(10^{-1}) = \{\emptyset, [\dots\{3\}], \{1, 2\}\},$$

$$\mathfrak{U}^{\text{q-opt}}(10^{-2}) = \mathfrak{U}^{\text{opt}}(10^{-2}) = \{\emptyset, [\dots\{4\}], \{1, 2\}, \{1, 3\}\},$$

$$\mathfrak{U}^{\text{PW}}(10^{-2}) = \{\emptyset, [\dots\{9\}], [\dots\{1, 7\}], \{2, 3\}, \{1, 2, 3\}\},$$

$$\mathfrak{U}^{\text{q-opt}}(10^{-3}) = \{\emptyset, [\dots\{12\}], [\dots\{1, 10\}], [\dots\{2, 5\}], \{1, 2, 3\}\},$$

$$\mathfrak{U}^{\text{opt}}(10^{-3}) = \{\emptyset, [\dots\{11\}], [\dots\{1, 9\}], \{2, 3\}, \{2, 4\}, \{1, 2, 3\}, \{1, 2, 4\}\},$$

$$\mathfrak{U}^{\text{PW}}(10^{-3}) = \{\emptyset, [\dots\{26\}], [\dots\{1, 21\}], [\dots\{2, 10\}], [\dots\{3, 7\}], \{4, 5\}, [\dots\{1, 2, 9\}], [\dots\{1, 3, 6\}]\}.$$

Case of $p = 2$ and $a = 2$

$$\mathfrak{U}^{\text{opt}}(10^{-1}) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\} \quad \text{and} \quad \mathfrak{U}^{\text{q-opt}}(10^{-1}) = \{\emptyset, [\dots\{4\}], \{1, 2\}\},$$

$$\mathfrak{U}^{\text{PW}}(10^{-1}) = \{\emptyset, [\dots\{8\}], [\dots\{1, 6\}], \{2, 3\}\}$$

$$\mathfrak{U}^{\text{q-opt}}(10^{-2}) = \{\emptyset, [\dots\{18\}], [\dots\{1, 10\}], [\dots\{2, 5\}], \{1, 2, 3\}\},$$

$$\mathfrak{U}^{\text{opt}}(10^{-2}) = \{\emptyset, [\dots\{14\}], [\dots\{1, 11\}], [\dots\{2, 5\}], [\dots\{1, 2, 4\}]\},$$

$$\begin{aligned} \mathfrak{U}^{\text{PW}}(10^{-2}) = & \{\emptyset, [\dots\{54\}], [\dots\{1, 41\}], [\dots\{2, 20\}], [\dots\{3, 13\}], [\dots\{4, 10\}], [\dots\{5, 8\}], [\dots\{1, 2, 15\}], \\ & [\dots\{1, 3, 10\}], [\dots\{1, 4, 7\}], \{1, 5, 6\}, \{2, 3, 4\}, \{2, 3, 5\}\}. \end{aligned}$$

Case $p = \infty$ and $a = 4$

$$\mathfrak{U}^{\text{q-opt}}(10^{-1}) = \mathfrak{U}^{\text{opt}}(10^{-1}) = \{\emptyset, \{1\}\} \quad \text{and} \quad \mathfrak{U}^{\text{PW}}(10^{-1}) = \{\emptyset, [\dots\{4\}], \{1, 2\}, \{1, 3\}\}.$$

$$\mathfrak{U}^{\text{q-opt}}(10^{-2}) = \mathfrak{U}^{\text{opt}}(10^{-2}) = \{\emptyset, [\dots\{3\}], \{1, 2\}\},$$

$$\mathfrak{U}^{\text{PW}}(10^{-2}) = \{\emptyset, [\dots\{10\}], [\dots\{1, 8\}], \{2, 3\}, \{2, 4\}, \{1, 2, 3\}\},$$

$$\mathfrak{U}^{\text{q-opt}}(10^{-3}) = \{\emptyset, [\dots\{8\}], [\dots\{1, 7\}]\},$$

$$\mathfrak{U}^{\text{opt}}(10^{-3}) = \{\emptyset, [\dots\{8\}], [\dots\{1, 6\}], \{2, 3\}\}$$

$$\mathfrak{U}^{\text{PW}}(10^{-3}) = \{\emptyset, [\dots\{26\}], [\dots\{1, 22\}], [\dots\{2, 11\}], [\dots\{3, 7\}], \{4, 5\}, [\dots\{1, 2, 9\}], [\dots\{1, 3, 6\}]\}$$

Case $p = \infty$ and $a = 3$

$$\mathfrak{U}^{\text{q-opt}}(10^{-1}) = \mathfrak{U}^{\text{opt}}(10^{-1}) = \{\emptyset, \{1\}, \{2\}\},$$

$$\mathfrak{U}^{\text{PW}}(10^{-1}) = \{\emptyset, [\dots\{10\}], [\dots\{1, 8\}], \{2, 3\}, \{2, 4\}, \{1, 2, 3\}\},$$

$$\mathfrak{U}^{\text{q-opt}}(10^{-2}) = \mathfrak{U}^{\text{opt}}(10^{-2}) = \{\emptyset, [\dots\{8\}], [\dots\{1, 6\}], \{2, 3\}\},$$

$$\begin{aligned}
\mathfrak{U}^{\text{PW}}(10^{-2}) &= \{\emptyset, [\dots\{49\}], [\dots\{1, 39\}], [\dots\{2, 19\}], [\dots\{3, 13\}], [\dots\{4, 9\}], \{5, 6\}, \\
&\quad \{5, 7\}, [\dots\{1, 2, 15\}], [\dots\{1, 3, 10\}], [\dots\{1, 4, 7\}], \{1, 5, 6\}, \{2, 3, 4\}, \{2, 3, 5\}, \{1, 2, 3, 4\}\}, \\
\mathfrak{U}^{\text{q-opt}}(10^{-3}) &= \{\emptyset, [\dots\{36\}], [\dots\{1, 29\}], [\dots\{2, 14\}], [\dots\{3, 9\}], [\dots\{4, 7\}], [\dots\{1, 2, 8\}]\}, \\
\mathfrak{U}^{\text{opt}}(10^{-3}) &= \{\emptyset, [\dots\{31\}], [\dots\{1, 25\}], [\dots\{2, 12\}], [\dots\{3, 8\}], [\dots\{4, 6\}], [\dots\{1, 2, 9\}], [\dots\{1, 3, 6\}]\},
\end{aligned}$$

$$\begin{aligned}
\mathfrak{U}^{\text{PW}}(10^{-3}) &= \{\emptyset, [\dots\{208\}], [\dots\{1, 165\}], [\dots\{2, 82\}], [\dots\{3, 55\}], [\dots\{4, 41\}], [\dots\{5, 33\}], [\dots\{6, 27\}], \\
&\quad [\dots\{7, 23\}], [\dots\{8, 20\}], [\dots\{9, 18\}], [\dots\{10, 16\}], [\dots\{11, 15\}], \{12, 13\}, [\dots\{1, 2, 65\}], [\dots\{1, 3, 43\}], \\
&\quad [\dots\{1, 4, 32\}], [\dots\{1, 5, 26\}], [\dots\{1, 6, 21\}], [\dots\{1, 7, 18\}], [\dots\{1, 8, 16\}], [\dots\{1, 9, 14\}], [\dots\{1, 10, 13\}], \\
&\quad [\dots\{2, 3, 21\}], [\dots\{2, 4, 16\}], [\dots\{2, 5, 13\}], [\dots\{2, 6, 10\}], \{2, 7, 8\}, \{2, 7, 9\}, [\dots\{3, 4, 10\}], [\dots\{3, 5, 8\}], \\
&\quad \{3, 6, 7\}, \{4, 5, 6\}, [\dots\{1, 2, 3, 17\}], [\dots\{1, 2, 4, 13\}], [\dots\{1, 2, 5, 10\}], \{1, 2, 6, 7\}, \{1, 2, 6, 8\}, \\
&\quad [\dots\{1, 3, 4, 8\}], \{1, 3, 5, 6\}\}
\end{aligned}$$

Case $p = \infty$ and $a = 2$

$$\begin{aligned}
\mathfrak{U}^{\text{q-opt}}(10^{-1}) &= \{\emptyset, [\dots\{22\}], [\dots\{1, 15\}], \{2, 3\}\} \\
\mathfrak{U}^{\text{opt}}(10^{-1}) &= \{\emptyset, [\dots\{16\}], [\dots\{1, 12\}], [\dots\{2, 5\}], [\dots\{1, 2, 4\}]\},
\end{aligned}$$

$$\begin{aligned}
\mathfrak{U}^{\text{PW}}(10^{-1}) &= \{\emptyset, [\dots\{511\}], [\dots\{1, 361\}], [\dots\{2, 180\}], [\dots\{3, 120\}], [\dots\{4, 90\}], [\dots\{5, 72\}], [\dots\{6, 60\}], \\
&\quad [\dots\{7, 51\}], [\dots\{8, 45\}], [\dots\{9, 40\}], [\dots\{10, 36\}], [\dots\{11, 36\}], [\dots\{11, 32\}], [\dots\{12, 30\}], [\dots\{13, 27\}], \\
&\quad [\dots\{14, 25\}], [\dots\{15, 24\}], [\dots\{16, 22\}], [\dots\{17, 21\}], [\dots\{18, 20\}], [\dots\{1, 2, 127\}], [\dots\{1, 3, 85\}], \\
&\quad [\dots\{1, 4, 63\}], [\dots\{1, 5, 51\}], [\dots\{1, 6, 42\}], [\dots\{1, 7, 36\}], [\dots\{1, 8, 31\}], [\dots\{1, 9, 28\}], [\dots\{1, 10, 25\}], \\
&\quad [\dots\{1, 11, 23\}], [\dots\{1, 12, 21\}], [\dots\{1, 13, 19\}], [\dots\{1, 14, 18\}], [\dots\{1, 15, 17\}], [\dots\{2, 3, 42\}], \\
&\quad [\dots\{2, 4, 31\}], [\dots\{2, 5, 25\}], [\dots\{2, 6, 21\}], [\dots\{2, 7, 18\}], [\dots\{2, 8, 15\}], [\dots\{2, 9, 14\}], [\dots\{2, 10, 12\}], \\
&\quad [\dots\{3, 4, 21\}], [\dots\{3, 5, 17\}], [\dots\{3, 6, 14\}], [\dots\{3, 7, 12\}], [\dots\{3, 8, 10\}], [\dots\{4, 5, 12\}], [\dots\{4, 6, 10\}], \\
&\quad [\dots\{4, 7, 9\}], [\dots\{5, 6, 8\}], [\dots\{1, 2, 3, 30\}], [\dots\{1, 2, 4, 22\}], [\dots\{1, 2, 5, 18\}], [\dots\{1, 2, 6, 15\}], \\
&\quad [\dots\{1, 2, 7, 12\}], [\dots\{1, 2, 8, 11\}], \{1, 2, 9, 10\}, [\dots\{1, 3, 4, 15\}], [\dots\{1, 3, 5, 12\}], [\dots\{1, 3, 6, 10\}], \\
&\quad \{1, 3, 7, 8\}, [\dots\{1, 4, 5, 9\}], \{1, 4, 6, 7\}, [\dots\{2, 3, 4, 7\}], \{2, 3, 5, 6\}, \{1, 2, 3, 4, 5\}\}.
\end{aligned}$$

References

- [1] R. E. Caflisch, W. Morokoff, and A. B. Owen: Valuation of mortgage backed securities using Brownian bridges to reduce the effective dimension. *Journal of Computational Finance* 1: 27–46, 1997.
- [2] M. Hefter, K. Ritter, and G. W. Wasilkowski, On equivalence of weighted anchored and ANOVA spaces of functions with mixed smoothness of order one in L_1 or L_∞ norm, *J. Complexity* **32** (2016), 1-19.
- [3] P. Kritzer, F. Pillichshammer, and G. W. Wasilkowski, Very low truncation dimension for high dimensional integration under modest error demand, *J. Complexity* **35** (2016), 63-85.

- [4] F. Y. Kuo, I. H. Sloan, G. W. Wasilkowski, and H. Woźniakowski, Liberating the dimension, *J. Complexity* **26** (2010), 422–454.
- [5] R. Liu and A. Owen: Estimating mean dimensionality of analysis of variance decompositions. *J. Amer. Statist. Assoc.* 101: 712–721, 2006.
- [6] A. B. Owen: Effective dimension for weighted function spaces. Technical Report, Stanford University, 2012 and in revised version 2014. available under <http://statweb.stanford.edu/~owen/reports/effdim-periodic.pdf>
- [7] L. Plaskota and G. W. Wasilkowski, Tractability of infinite-dimensional integration in the worst case and randomized settings, *J. Complexity* **27** (2011), 505–518.
- [8] I. H. Sloan and H. Woźniakowski, When are Quasi-Monte Carlo algorithms efficient for high dimensional integrals?, *J. Complexity* **14** (1998), 1–33.
- [9] X. Wang and K.-T. Fang: Effective dimension and quasi-Monte Carlo integration. *J. Complexity* 19: 101–124, 2003.

Authors’ Addresses:

A. D. Gilbert
 School of Mathematics and Statistics
 The University of New South Wales
 Sydney, NSW 2052, Australia
 E-mail: alexander.gilbert@student.unsw.edu.au

G. W. Wasilkowski
 Department of Computer Science
 University of Kentucky
 301 David Marksbury Building
 Lexington, KY 40506, USA
 E-mail: greg@cs.uky.edu

Funding: The author A. D. Gilbert is grateful for the financial support received from the Australian Research Council (DP150101770) and the University of New South Wales.